AD_____

Award Number: W81XWH-08-1-0532

TITLE: Structure and Function of the Splice Variants of TMPRSS2-ERG, a Prevalent Genomic Alteration in Prostate Cancer

PRINCIPAL INVESTIGATOR: Shiv Srivastava, Ph.D.

CONTRACTING ORGANIZATION: Henry M. Jackson Foundation
Rockville, MD 20852

REPORT DATE: September 2009

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

X Approved for public release; distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE:** Annual	**3. DATES COVERED** *(From - To)*
Sept 1, 2009		Aug 15, 2008-Aug 14, 2009

4. TITLE AND SUBTITLE: Structure and Function of the Splice variants of TMPRSS2-ERG, a Prevalent Genomic Alteration in Prostate Cancer

5a. CONTRACT NUMBER: W81XWH-08-1-0532

5b. GRANT NUMBER: PC073614

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S): Shiv Srivastava;

Email: ssrivastava@cpdr.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Henry M. Jackson Foundation
Headquarters, 1401 Rockville Pike,
CPDR, Rockville, MD 20852

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research
Materiel Command
Fort Detrick, MD 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release: Distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Characterization of major *TMPRSS2-ERG* splice variants in prostate cancer specimens for the first time provides information on the full length protein coding sequences of *TMPRSS2-ERG* fusion transcripts which is urgently needed in developing more precise bio-marker and therapeutic strategies. Unexpected discovery on predominance of type II splice variants e.g., *TMPRSS2-ERG8* and identification of novel splice variants: *TMPRSS2-EPC1 and TMPRSS2-EPC2* represents the unique outcome of this grant effort focusing on unbiased evaluation of fusion transcripts in prostate cancer. Definition of the relative abundance of the *TMPRSS2-ERG* splice variants in prostate cancer provide novel information with respect to type II *TMPRSS2-ERG* transcripts and prognostic value of measuring the ratio of type II and type I transcripts. These data suggest that over all biologic effects of *TMPRSS2-ERG* fusion may be a reflection of potential interaction of these splice variants.

15. SUBJECT TERMS None provided.

16. SECURITY CLASSIFICATION OF: U			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					USAMRMC
a. REPORT: U	**b. ABSTRACT:** U	**c. THIS PAGE:**	UU	30	**19b. TELEPHONE NUMBER** *(include area code)*

Table of Contents

1 INTRODUCTION

With in the last three year, accumulating data from several groups including our laboratory have confirmed that *TMPRSS2-ERG* fusion is predominant among other *TMPRSS2-ETS* factor fusions (Kumar-Sinha et al., 2008). These data underscore that frequent *ERG* proto-oncogene overexpression (Petrovics et al., 2005) due to androgen regulated *TMPRSS2* promoter (Lin et al., 1999; Nelson et al., 1999; Tomlins et al., 2005)in prostate cancer (CaP) cells represents a high frequency (~60-70%) and potentially causal oncogenic alteration in CaP (Carver et al., 2009; King et al., 2009; Klezovitch et al., 2008; Sun et al., 2008; Tomlins et al., 2008). *ETS* family encoded proteins show a wide variety of expression patterns in human tissues (Turner and Watson, 2008). Similar to the other members of *ETS* family, *ERG* plays a central role in mediating mitogenic signals through major cellular pathways, including the mitogen activator protein kinases (MAPK) (Hart et al., 1995). *ERG* gene consists of 17 exons that spans about 300 kb, and is transcribed to generate nine alternate splice forms (Owczarek et al., 2004). At least five of the ERG splice variants are translated into proteins, ERG-1 (p41), ERG-2 (p52), ERG-3 (p55), ERG-4 (p49) and ERG-5 (p38, predicted MW 46 kDa) by a combination of alternative mRNA splicing and/or use of alternative polyadenlyation sites. These proteins exert their functions as transcriptional regulators (Turner and Watson, 2008). ERG proteins consist of a conserved PNT (or Pointed) domain at the N-terminal and DNA-binding *ETS* domain at the C-terminal end. *ERG* encoded proteins bind to the core DNA sequence 5'-GGA(A/T) - 3' on the promoter region and regulate the downstream target gene expression, either as homo or heterodimers with other ETS proteins. Several reports indicate that *ERG* functions as a proto-oncogene with transforming activity (Oikawa and Yamada, 2003; Rainis et al., 2005; Rao et al., 1987; Sementchenko et al., 1998). Our study integrating complementary data from GeneChips, quantitative PCR and RNA *in situ* hybridization in a relatively large cohort of CaP patients highlighted that *ERG* is over-expressed in about 70% of prostate cancer specimens (Hu et al., 2008; Petrovics et al., 2005). In light of *TMPRSS-ERG* gene fusion as a primary mechanism of the *ERG* overexpression in CaP cells, important questions with respect to structure, expression and functions of the *ERG* remains to be defined in CaP cells. The different types of fusion transcripts (full length) and the encoded protein products expressed in CaP cells as a result of the *TMPRSS-ERG* fusion have to be defined. Most of the published reports thus far have focused on the *TMPRSS2-ERG* fusion junctions at the 5' end with a focus on PCR products of already annotated sequences of *ERG* where as most of the splicing events occur 3' to fusion (Wang et al., 2008). Much remains to be defined with respect to relative abundance or functions of specific *ERG* splice variants in the context of CaP cells.

This proposal is built on the hypothesis that specific *ERG* splice forms in *TMPRSS2-ERG* fusion configuration are selectively expressed in CaP cells and are functionally relevant in CaP. Although, it has been suggested that some fusion variants are associated with poor prognosis, careful evaluation of individual full length *TMPRSS2-ERG* transcripts is necessary to understand the biologic functions of *TMPRSS2-ERG* transcripts. Towards defining full length *TMPRSS2-ERG* transcripts including the specific splice variants of *ERG* in CaP cells we have generated a cDNA library from tumor specimens of CaP patients. Delineation of structure and functions of full length *TMPRSS2-ERG* transcripts in CaP will lead to a major advancement in understanding the role of *ERG* in prostate tumor biology.

2 BODY

The progress reported here is built on the development of novel preliminary data reported in original proposal.

Task #1: Characterization of full length sequences of *TMPRSS2-ERG* transcripts in prostate cancer.

Month 0-12: *Overall aim is to perform innovative analyses of CPDR-CaP- library (at least 1,000,000 pfus) that would facilitate the identification of relatively common TMPRSS2-ERG splice and fusion variants in human CaP.*

- *CPDR CaP library will screened with ERG cDNA probe and by TMPRSS2 probe to enrich for TMPRSS2-ERG fusion cDNA clones.*
- *3'-RACE will be performed to identify the authentic 3' UTRs from the mRNA pool from CaP patient specimens. Using phage excision strategy, we will use to generate plasmids (cDNA) clones.*
- *DNA sequence of the clones will be determined and the sequences will be evaluated for open reading frames to identify TMPRSS2-ERG fusion variants and other ERG splice variants.*

Accomplished:
The CPDR CaP library was screened with *ERG* cDNA probe and by *TMPRSS2* probe to enrich for *TMPRSS2-ERG* fusion cDNA clones. The identified clones contained 3' polyadenylation signals that clearly defined the 3' UTRs. The 5' transcription start sites were precisely defined by 5'-RACE method.Clones were sequenced and analyzed for ERG coding reading frames. From the phage DNA sequences plasmids (cDNA) clones were generated by using phage excision strategy.

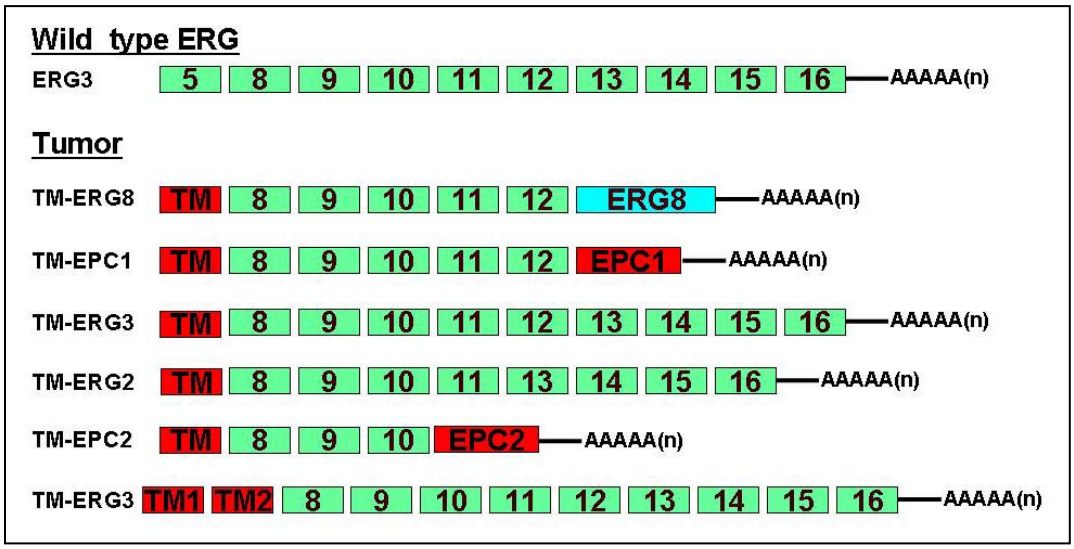


Figure 1. *ERG* splice variants in prostate cancer patients. As proposed we obtained representative number of full-length *TMPRSS2-ERG* clones of all major types.

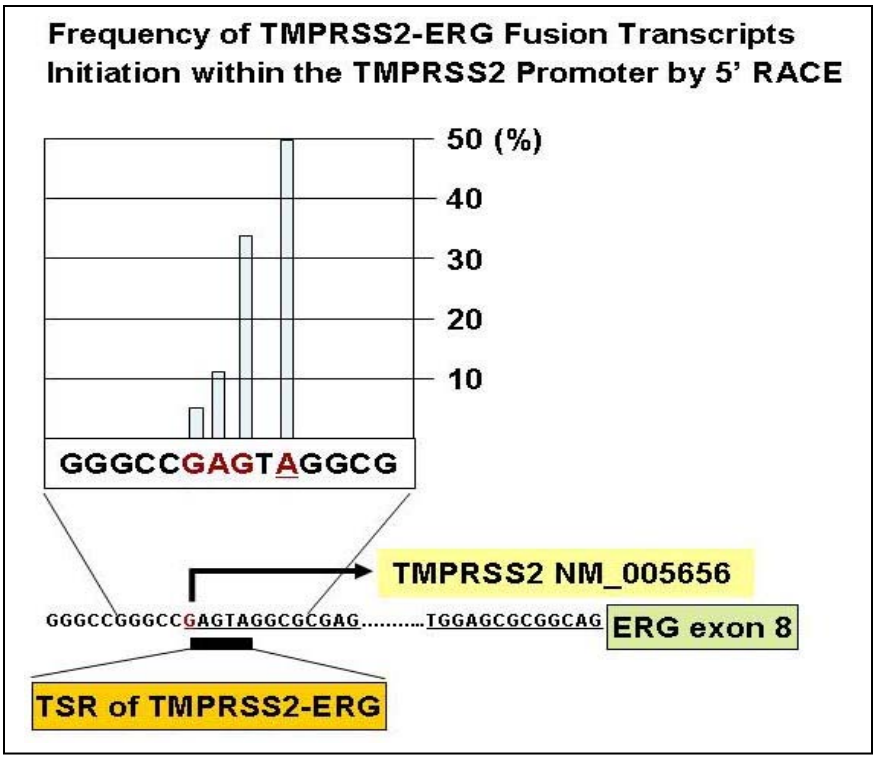


Figure 2. Frequency of transcription initiation of TMPRSS2-ERG fusion transcripts was defined by 5' RACE method from twenty clones. Transcription initiation site of the wild type (NM_005656) TMPRSS2 transcript relative to the transcription start region (TSR) of TMPRSS2-ERG fusion transcripts and the TMPRSS2-to-ERG exon 8 junction are also shown.

Novel Findings:

- Extensive analysis of clones revealed tumor specific novel sequences with TMPRSS2 gene fusions such as TMPRSS2-ERG8, TMPRSS2-EPC1 and TMPRESS2-EPC2.
- Translational capabilities of these identified clones were analyzed by subcloning them in expression vectors and transfected into HEK293 and LNCaP cells (Figure 3.)

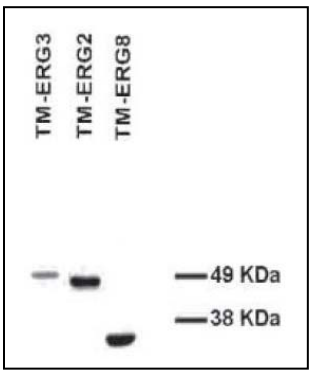


Figure 3. Detection of full-length ERG proteins, products of TMPRSS2-ERG3, TMPRSS2-ERG2 and TMPRSS2-ERG8 coding sequences by immunoblot assay.

<u>Of note, characterization of these clones for the first time provides information on the full length protein coding sequences of TMPRSS2-ERG fusion transcripts which is urgently needed in developing more precise bio-marker and therapeutic strategies.</u>

- Antibodies to ERG proteins were developed and characterized to detect type I (Figure 4A-B) and type II (ERG8) (Figure 4C).

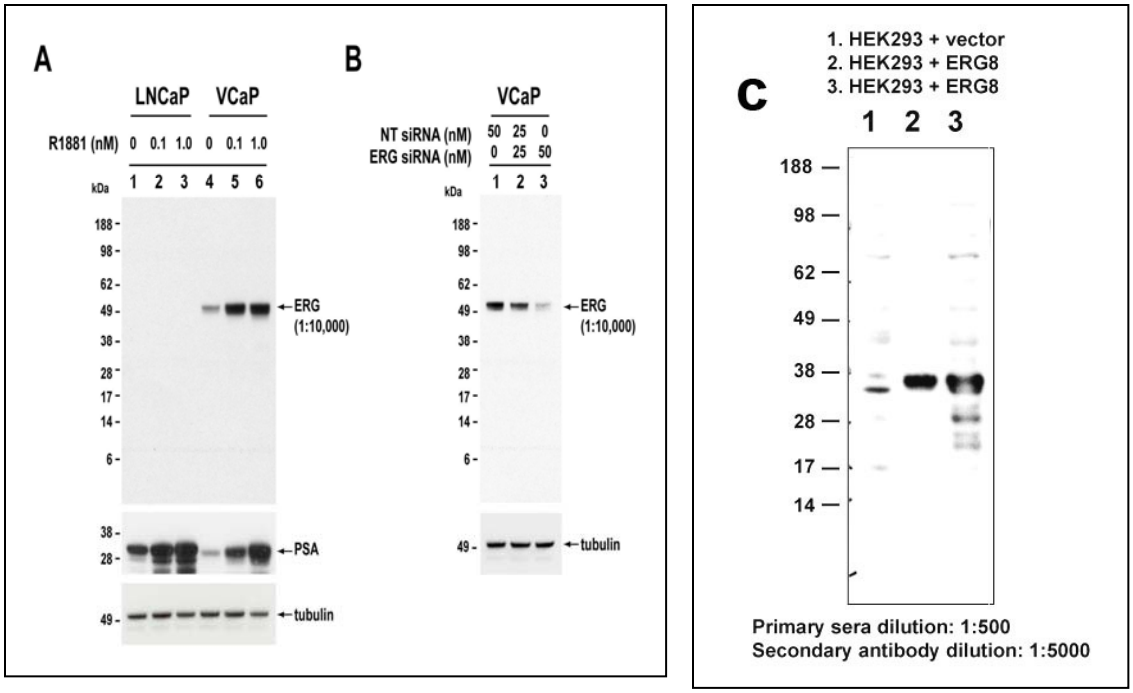


Figure 4. Detection of Type I ERG protein by ERG antibody. A, Endogenous ERG protein expression in TMPRSS2-ERG harboring VCaP cells in response to androgen treatment (R1881). B, Knockdown of ERG protein expression by specific ERG siRNA. C, Immunodetection of ERG8 protein by a polyclonal anti-ERG8 antibody raised against an ERG8-specific C-terminal peptide.

Task #2: Quantitative evaluations of selected *TMPRSS2-ERG* variants in prostate cancer specimens and prognostic features.

Month 6-24: *From the CPDR prostate cancer mRNA bank transcripts of the ERG locus will be defined by quantitative PCR technique.*

- *Towards this task, we will perform parallel quantitative analyses up to 6 most abundant TMPRSS2-ERG variants in laser capture micro dissected tumor and matching benign epithelial cells from 150 patients representing primary prostate cancer specimens.*
- *To evaluate association of various TMPRSS2-ERG variants with clinico-pathologic parameters, quantitative expression will be assessed for patient age, race, pre-treatment PSA level, Gleason score, CaP family history, tumor stage, surgical margin status, seminal vesicle invasion, PSA recurrence, bone metastasis, nuclear grade, differentiation, PSA doubling time and prostate cancer associated death.*
- *The associations of TMPRSS2-ERG transcripts with FISH defined genomic deletions associating with poor prognosis will be assessed*

- *We will assess the relative abundance of normal ERG splice variants (based on literature information and aim 1) in pooled normal prostate RNA and in LCM-benign epithelial cells of 25 patient specimens (CPDR tissue bank).*

Accomplished:

Quantitative features and prognostic potential of selected *TMPRSS2-ERG* variants in prostate cancer specimen were determined in tumor and matching benign specimens from 122 patients. Transcripts of the ERG locus were defined by quantitative PCR technique by using the CPDR prostate cancer mRNA bank. Towards this task, we performed the parallel quantitative analyses of TM-ERG8, TM-EPC1, TM-ERG1, TM-ERG2 and TM-ERG3 splice variants in laser capture micro dissected tumor and matching benign epithelial cells from 122 patients representing primary prostate cancer specimens (Figure 5). By evaluating quantitative expression features in tumors from 122 patients the data reached statistical significance in the clinico-pathologic correlation analysis.

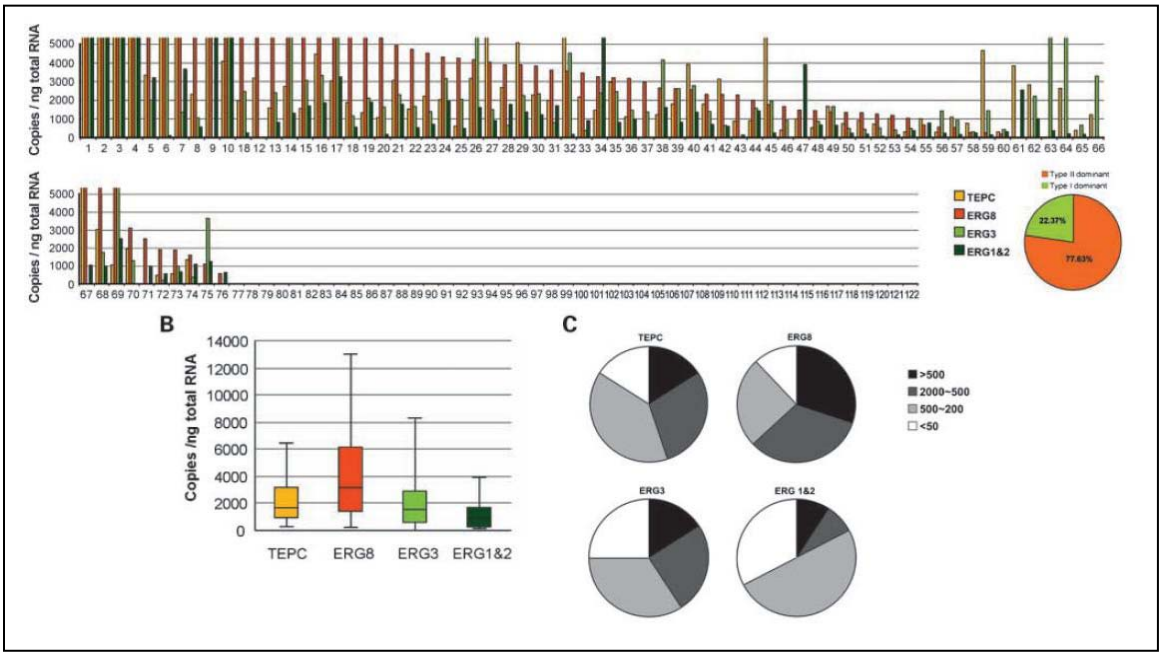


Figure 5. Quantitative expression of *ERG* splice variants in prostate cancer patients. *A*, Quantitative expression of *ERG* splice forms *ERG8, TEPC, ERG3* and *ERG 1&2* (represented by columns with different colors) were determined in microdissected tumor cells of CaP patients (N=122). The graph depicts relative expression levels (normalized to *GAPDH*) in patients with (upper panel, N=66) and with no detectable (lower panel, N=56) *TMPRSS2-ERG* fusion A transcript. In the pie chart the % of CaP patients are represented either with higher (~77%) or with lower (~23%) expression level of type II (*ERG*8 and *EPC*1) than type I *ERG* (*ERG*1-3) transcripts. *B,* Relative abundance of various *ERG* splice forms (*ERG8, TEPC, ERG3* and *ERG 1-2*) is depicted by boxplots representing the copy numbers determined in 76 CaP patients overexpressing *ERG*. C, Pie charts illustrate the distribution of CaP patients with various expression levels (copy number/ng total RNA) of *ERG8, TEPC, ERG3* and *ERG 1&2* splice forms in CaP cells. <u>These data for the first time carefully define the relative

abundance of the *TMPRSS2-ERG* splice variants in prostate cancer and provide novel information with respect to type II *TMPRSS2-ERG* transcripts and prognostic value of measuring the ratio of type II and type I transcripts.

Towards defining the relative abundance and the ratios of the full-length fusion transcripts we evaluated the association of various *TMPRSS2-ERG* variants with clinico-pathologic parameters, quantitative expression was assessed for patient age, race, pre-treatment PSA level, Gleason score, CaP family history, tumor stage, surgical margin status, seminal vesicle invasion, PSA recurrence, bone metastasis, nuclear grade, differentiation, PSA doubling time and prostate cancer associated death. This study addressed the potential of splice variants with important biological and clinical relevance. A remarkable finding of this study is that an increased ration of type I *ERG* (*ERG*1-3) over type II (*ERG*8 and *EPC*1) splice variants tightly correlates with higher pathological Gleason sum and poor LCM differentiation (Figure 6).

Clinico-pathological characteristics	ERG splice variant expression			Type I/ type II ratio		
	No (46)	Yes (76)	P value**	N (76)	median	P value*
Race			**0.0345**			0.1142
Caucasian	29 (33%)	59 (67%)		59	0.51	
African American	15 (56%)	12 (44%)		12	0.33	
Family history			0.6312			0.3259
No	28 (38%)	45 (62%)		45	0.52	
Yes	10 (33%)	20 (67%)		20	0.51	
Pathological T stage			0.5318			0.2541
pT2	14 (36%)	25 (64%)		25	0.44	
pT3	29 (42%)	40 (58%)		40	0.51	
Pathological Gleason sum			**0.0023**			**0.0323**
2 to 6	9 (24%)	29 (76%)		29	0.35	
7	20 (36%)	36 (64%)		36	0.46	
8 to 10	14 (70%)	6 (30%)		6	0.70	
LCM differentiation			**0.0058**			**0.0067**
Well	31 (32%)	66 (68%)		65	0.45	
Poorly	15 (62%)	9 (38%)		9	0.76	
Margin status			0.9436			**0.0032**
Negative	28 (38%)	45 (62%)		45	0.39	
Positive	16 (39%)	25 (61%)		25	0.57	
PSA recurrence			0.7312			**0.0456**
No	33 (37%)	56 (63%)		56	0.42	
Yes	11 (41%)	16 (59%)		16	0.61	

* one-sided test, $P<0.05$ was considered statistically significant.
** two-sided test, $P<0.05$ was considered statistically significant.

Figure 6. Type I/type II ratio (N=76): The ratio of *ERG* typeI/typeII splice variants in CaP cells is increased in patients with poor tumor cell differentiation and with PSA recurrence.

Towards the completion of Task#2 the following experiments will be performed in the next reporting period:

- Evaluation of the associations of *TMPRSS2-ERG* transcripts with FISH defined genomic deletions and their correlation with poor prognosis

- Evaluation of the relative abundance of normal *ERG* splice variants in pooled normal prostate RNA and in LCM-benign epithelial cells of 25 patient specimens (CPDR tissue bank).

Task #3: Defining the functional significance of specific splice variants of the cancerous ERG locus.

Month 6-36: In VCaP cell culture model system, we will assess whether *TMPRSS2-ERG* variants (type I and type II) contribute to the net oncogenic effect of *ERG*.

- Multiple inhibitory siRNA molecules will be designed to specifically target identified individual transcripts. Dose response and kinetic measures of inhibitory efficacy for each molecule will be monitored by measuring the expression of specific *TMPRSS2-ERG* transcript and protein targeted by the selected siRNA.
- Most effective siRNA molecules will be selected to evaluate the combined effect of multiple siRNA molecules and measure dose response and kinetic characteristics.
- Individual *TMPRSS2-ERG* variants (type I and type II) will be over expressed using *TMPRSS2* promoter driven expression vectors in telomerase immortalized normal prostate cancer cells, RC165 established in CPDR. Cancer biology related features (cell growth, soft agar colony formation, cell invasion and changes in cell cycle of cell) of the cells with knock down or ectopic expression specific *TMPRSS2-ERG* transcripts will be measured in response.
- Predicted protein products from type I and type II *TMPRSS2-ERG* transcripts will be tested for their ability and specificity to bind to target sequences in gel-shift experiments. The regulatory efficiency of type I and type II products will be tested using different known promoters (TGF beta type II receptor, MMP3 and collagen) using luciferase assay systems.

Towards Task#3 inhibitory siRNA molecules were designed and prepared to specifically target the abundant ERG8 (Type II) and ERG1-3 (Type I) transcripts. The dose and kinetic assessment of siRNA molecules are in progress. The regulatory efficiency of type I and type II products are being tested using different known promoters by using luciferase assay systems.

3 KEY RESEARCH ACCOMPLISHMENTS

Task #1 Characterization of full length sequences of *TMPRSS2-ERG* transcripts in prostate cancer was addressed.

- CPDR CaP library screening was performed with *ERG* cDNA probe and by *TMPRSS2* probe to enrich for *TMPRSS2-ERG* fusion cDNA clones.
- Full-length cDNA clones were obtained from the mRNA pool from CaP patient specimens with polyadenylation signals that defined the 3' untranslated sequences of clones. We generated plasmids (cDNA) clones of *TMPRSS2-ERG* splice variants by using phage excision strategy.
- DNA sequences of the clones were determined and were evaluated for open reading frames. We identify *TMPRSS2-ERG* splice variants from the clones. Novel cDNA sequences were identified by these experiments such as, TM-EPC1 and TM-EPC2.

Task #2 Quantitative assessment of TMPRSS2-ERG splice variants and the correlation analysis of splice variants with clinico-pathologic parameters have been evaluated.

- We performed parallel quantitative analyses of the five most abundant *TMPRSS2-ERG* splice variants in laser capture micro dissected tumor and matching benign epithelial cells from 122 patients representing primary prostate cancer specimens. The number of specimens examined here provided statistically significant information in subsequent studies.
- We evaluated the association of various *TMPRSS2-ERG* variants with clinico-pathologic parameters, such as, patient age, race, pre-treatment PSA level, Gleason score, CaP family history, tumor stage, surgical margin status, seminal vesicle invasion, PSA recurrence, bone metastasis, nuclear grade, differentiation, PSA doubling time and prostate cancer associated death. Significant correlation was found between increased ratio of type I *ERG* (*ERG*1-3) over type II (*ERG*8 and *EPC*1) splice variants and higher pathological Gleason sum and poor LCM differentiation

4 REPORTABLE OUTCOMES

Publication: Hu Y, Dobi A, Sreenath T, Cook C, Tadese AY, Ravindranath L, Cullen J, Furusato B, Chen Y, Thangapazham R, Mohamed A, Sun C, Sesterhenn IA, McLeod DG, Petrovics G, Srivastava S.: Delineation of *TMPRSS2-ERG* splice variants in prostate cancer. *Clinical Cancer Research* 14, 4719-4725 (2008).

International Patent application: Prostate cancer-specific alterations in ERG gene expression and detection and treatment methods based on those alterations. Pub. No.: WO/2008/063769 International Application No.: PCT/US2007/080826, 2008.

Press release: Researchers provide new insights into common alterations of ERG oncogene in prostate cancer. Uniformed Services University, Bethesda, Maryland, October 8, 2008 (www.usuhs.mil)

Podium presentation: 9th Asian Congress of Urology New Delhi, India, Oct 4, 2008. Shiv Srivastava: Bio-marker and Therapeutic Target Potential of Frequent Gene Alterations in Prostate Cancer: State-of-the-art lecture

Poster presentations:
1) Kimbrough Urological Seminar-Washington DC, Jan 11-16, 2009. **Best Poster Award**. Albert Dobi, Chen Sun, Ying Hu, Rajesh Thangapazham, Ahmed Mohamed, CPT Eric J. Whitman, CPT Dorotha Hawksworth, Bungo Furusato, Shyh-Han Tan, Atekelt Y. Tadase, Lakshmi Ravindranath, Jennifer Cullen, Yongmei Chen, Gyorgy Petrovics, Taduru Sreenath, Isabell A. Sestrehenn, COL David G. McLeod, Shiv Srivastava: Splice Variants and Functions of *TMPRSS2-ERG* Fusion, A Common Genomic Alteration in Prostate Cancer

2) Society for Basic Urology Research, Phoenix, AZ, Nov 20-23, 2008. Shyh-Han Tan, Ying Hu, Albert Dobi, Taduru Sreenath, Christopher Cook, Atekelt Y. Tadase, Lakshmi Ravindranath, Jennifer Cullen, Bungo Furusato, Yongmei Chen, Rajesh Thangapazham, Ahmed Mohamed, Chen Sun, Isabell Sesterhenn, David G. McCleod and Shiv Srivastava: *TMPRSS2-ERG* Splice Variants In Prostate Cancer

3) Advances in Prostate cancer (AACR) San Diego, CA Jan 21-24, 2009. Taduru L. Sreenath, Ying Hu, Albert Dobi, Atekelt Tadese, Shyh-Han Tan, Lakshmi Ravindranath, Bungo Furusato, Youngmei Chen, Rajesh Thangapazham, Amina Ali, Ahmed Mohamed, Chen Sun, Isabell Sesterhenn, David G. McLeod, Gyorgy Petrovics, Shiv Srivastava: Expression and Sub-cellular Localization of Predominant TMPRSS2-ERG Transcripts with and without ETS domain in Prostate Cancer

4) American Association for Cancer Research, Denver CO, April 18-22, 2009
Ying Hu, Albert Dobi, Taduru Sreenath, Atekelt Tadese, Shyh-Han Tan, Lakshmi Ravindranath, Bungo Furusato, Yongmei Chen, Rajesh Thangapazham, Amina Ali, Ahmed Mohamed, Chen Sun, Isabell Sesterhenn, David McLeod, Gyorgy Petrovics, Shiv Srivastava: Predominant TMPRSS2-ERG Transcripts in Prostate Cancer

5) Uniformed Services University of the Health Sciences Research Week, May 11-13, 2009. Taduru Sreenath, Ying Hu, Albert Dobi, Atekelt Y Tadese, Shyh-Han Tan, Lakshmi Ravindranath, Bungo Furusato, Yongmei Chen, Rajesh Thangapazham, Amina Ali, Ahmed Mohamed, Chen Sun, Isabell Sesterhenn, David G McLeod, Gyorgy Petrovics, Shiv Srivastava: Differential Expressions and Sub-cellular Localization of TMPRSS2-ERG Splice Variants in Prostate Cancer

5 CONCLUSIONS:

1.Characterization of major *TMPRSS2-ERG* splice variants in prostate cancer specimens for the first time provides information on the full length protein coding sequences of *TMPRSS2-ERG* fusion transcripts which is urgently needed in developing more precise bio-marker and therapeutic strategies.

2. Unexpected discovery on predominance of type II splice variants e.g., *TMPRSS2-ERG8* and identification of novel splice variants: *TMPRSS2-EPC1 and TMPRSS2-EPC2* represent the unique outcome of this grant effort focusing on unbiased evaluation of fusion transcripts in prostate cancer.

3. Definition of the relative abundance of the *TMPRSS2-ERG* splice variants in prostate cancer provide novel information with respect to type II *TMPRSS2-ERG* transcripts and prognostic value of measuring the ratio of type II and type I transcripts. These data suggest that over all biologic effects of *TMPRSS2-ERG* fusion may be a refelction of potential interaction of these splice variants.

6 REFERENCES

Carver, B.S., Tran, J., Gopalan, A., Chen, Z., Shaikh, S., Carracedo, A., Alimonti, A., Nardella, C., Varmeh, S., Scardino, P.T., *et al.* (2009). Aberrant ERG expression cooperates with loss of PTEN to promote cancer progression in the prostate. Nat Genet *41*, 619-624.

Hart, A.H., Corrick, C.M., Tymms, M.J., Hertzog, P.J., and Kola, I. (1995). Human ERG is a proto-oncogene with mitogenic and transforming activity. Oncogene *10*, 1423-1430.

Hu, Y., Dobi, A., Sreenath, T., Cook, C., Tadase, A.Y., Ravindranath, L., Cullen, J., Furusato, B., Chen, Y., Thangapazham, R.L., *et al.* (2008). Delineation of TMPRSS2-ERG splice variants in prostate cancer. Clin Cancer Res *14*, 4719-4725.

King, J.C., Xu, J., Wongvipat, J., Hieronymus, H., Carver, B.S., Leung, D.H., Taylor, B.S., Sander, C., Cardiff, R.D., Couto, S.S., *et al.* (2009). Cooperativity of TMPRSS2-ERG with PI3-kinase pathway activation in prostate oncogenesis. Nat Genet *41*, 524-526.

Klezovitch, O., Risk, M., Coleman, I., Lucas, J.M., Null, M., True, L.D., Nelson, P.S., and Vasioukhin, V. (2008). A causal role for ERG in neoplastic transformation of prostate epithelium. Proc Natl Acad Sci U S A *105*, 2105-2110.

Kumar-Sinha, C., Tomlins, S.A., and Chinnaiyan, A.M. (2008). Recurrent gene fusions in prostate cancer. Nat Rev Cancer *8*, 497-511.

Lin, B., Ferguson, C., White, J.T., Wang, S., Vessella, R., True, L.D., Hood, L., and Nelson, P.S. (1999). Prostate-localized and androgen-regulated expression of the membrane-bound serine protease TMPRSS2. Cancer Res *59*, 4180-4184.

Nelson, P.S., Gan, L., Ferguson, C., Moss, P., Gelinas, R., Hood, L., and Wang, K. (1999). Molecular cloning and characterization of prostase, an androgen-regulated serine protease with prostate-restricted expression. Proc Natl Acad Sci U S A *96*, 3114-3119.

Oikawa, T., and Yamada, T. (2003). Molecular biology of the Ets family of transcription factors. Gene *303*, 11-34.

Owczarek, C.M., Portbury, K.J., Hardy, M.P., O'Leary, D.A., Kudoh, J., Shibuya, K., Shimizu, N., Kola, I., and Hertzog, P.J. (2004). Detailed mapping of the ERG-ETS2 interval of human chromosome 21 and comparison with the region of conserved synteny on mouse chromosome 16. Gene *324*, 65-77.

Petrovics, G., Liu, A., Shaheduzzaman, S., Furusato, B., Sun, C., Chen, Y., Nau, M., Ravindranath, L., Dobi, A., Srikantan, V., *et al.* (2005). Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome. Oncogene *24*, 3847-3852.

Rainis, L., Toki, T., Pimanda, J.E., Rosenthal, E., Machol, K., Strehl, S., Gottgens, B., Ito, E., and Izraeli, S. (2005). The proto-oncogene ERG in megakaryoblastic leukemias. Cancer Res *65*, 7596-7602.

Rao, V.N., Papas, T.S., and Reddy, E.S. (1987). erg, a human ets-related gene on chromosome 21: alternative splicing, polyadenylation, and translation. Science *237*, 635-639.

Sementchenko, V.I., Schweinfest, C.W., Papas, T.S., and Watson, D.K. (1998). ETS2 function is required to maintain the transformed state of human prostate cancer cells. Oncogene *17*, 2883-2888.

Sun, C., Dobi, A., Mohamed, A., Li, H., Thangapazham, R.L., Furusato, B., Shaheduzzaman, S., Tan, S.H., Vaidyanathan, G., Whitman, E., *et al.* (2008). TMPRSS2-ERG fusion, a common genomic alteration in prostate cancer activates C-MYC and abrogates prostate epithelial differentiation. Oncogene *27*, 5348-5353.

Tomlins, S.A., Laxman, B., Varambally, S., Cao, X., Yu, J., Helgeson, B.E., Cao, Q., Prensner, J.R., Rubin, M.A., Shah, R.B., *et al.* (2008). Role of the TMPRSS2-ERG gene fusion in prostate cancer. Neoplasia *10*, 177-188.

Tomlins, S.A., Rhodes, D.R., Perner, S., Dhanasekaran, S.M., Mehra, R., Sun, X.W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R., *et al.* (2005). Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science *310*, 644-648.

Turner, D.P., and Watson, D.K. (2008). ETS transcription factors: oncogenes and tumor suppressor genes as therapeutic targets for prostate cancer. Expert Rev Anticancer Ther *8*, 33-42.

Wang, J., Cai, Y., Yu, W., Ren, C., Spencer, D.M., and Ittmann, M. (2008). Pleiotropic biological activities of alternatively spliced TMPRSS2/ERG fusion gene transcripts. Cancer Res *68*, 8516-8524.

7 APPENDICES

Delineation of *TMPRSS2-ERG* Splice Variants in Prostate Cancer

Ying Hu,[1] Albert Dobi,[1,4] Taduru Sreenath,[1] Christopher Cook,[1] Atekelt Y. Tadase,[1] Lakshmi Ravindranath,[1] Jennifer Cullen,[1] Bungo Furusato,[1,2] Yongmei Chen,[1] Rajesh L. Thangapazham,[1] Ahmed Mohamed,[1] Chen Sun,[1] Isabell A. Sesterhenn,[2] David G. McLeod,[1,3,4] Gyorgy Petrovics,[1] and Shiv Srivastava[1,4]

Abstract

Purpose: The expression of the *ETS*-related gene (*ERG*) is low or undetectable in benign prostate epithelial cells. High prevalence of *ERG* overexpression in prostate cancer cells due to *TMPRSS2-ERG* fusions suggest for causal roles of ERG protein in the neoplastic process. *TMPRSS2-ERG* fusion junctions have been extensively studied in prostate cancer. However, virtually nothing is known about the nature of full-length transcripts and encoded proteins. This study focuses on qualitative and quantitative features of full-length *TMPRSS2-ERG* transcripts in prostate cancer.

Experimental Design: Full-length *TMPRSS2-ERG* transcripts were cloned and sequenced from a cDNA library generated from pooled RNA of six *TMPRSS2-ERG* fusion–positive prostate tumors. The encoded ERG proteins were analyzed in HEK293 cells. Copy numbers of *TMPRSS2-ERG* splice variants were determined by quantitative reverse transcription-PCR in laser capture microdissected prostate cancer cells.

Results: Two types of *TMPRSS2-ERG* cDNAs were identified: type I, which encodes full-length prototypical ERG protein (*ERG1, ERG2, ERG3*), and type II, encoding truncated ERG proteins lacking the ETS domain (*ERG8* and a new variant, *TEPC*). In microdissected prostate tumor cells from 122 patients, relative abundance of these variants was in the following order: *ERG8* > *TEPC* > *ERG 3* > *ERG1/2* with combined overexpression rate of 62.3% in prostate cancer. Increased ratio of type I over type II splice forms showed a trend of correlation with less favorable pathology and outcome.

Conclusions: Qualitative and quantitative features of specific *ERG* splice variants defined here promise to enhance the utility of *ERG* as a biomarker and therapeutic target in prostate cancer.

Molecular genetic evaluations of prostate cancer are defining mutational and expression alterations of critical oncogenes involved in disease onset and/or progression (reviewed in refs. 1–3). Discovery of prevalent chromosomal rearrangements/translocations leading to the activation of ETS transcription factors (predominantly *ERG*) through the androgen receptor–regulated *TMPRSS2* gene promoter underscore the critical roles of *ERG*-encoded protein in prostate cancer (4–7). Because *ERG* represents the majority of *TMPRSS2-ETS* factor alterations described thus far (6, 7), we have focused on the expression and regulation of *TMPRSS2-ERG* in prostate cancer. Oncogenic functions of ETS factors, including ERG, have also been implicated in diverse cancers (8).

Structure and function of *ERG*-encoded proteins remain to be defined in prostate cancer. *ERG* consists of 17 exons spanning about 300 kb and generates at least nine alternate splice forms, seven of them coding for protein products of varying sizes (9). These *ERG* splice variants have been primarily described in nonprostate tissues. Despite the large body of data on the *TMPRSS2-ERG* fusion junctions in prostate cancer (reviewed in refs. 6, 7), virtually nothing is known about the full-length *TMPRSS2-ERG* transcripts in prostate cancer, including the existence and relative abundance of specific splice variants. In this context, it is important to note that the cancer-associated splice variants of numerous genes, e.g., *androgen receptor*, *fibroblast growth factor receptor*, *survivin*, and *MDM2*, have functional implications (10, 11). Thus, characterization of full-length *TMPRSS2-ERG* transcripts is essential to better understand *ERG* function(s) in prostate cancer and to further enhance its utility as biomarker and therapeutic target.

Authors' Affiliations: [1]Center for Prostate Disease Research, Department of Surgery, Uniformed Services University of the Health Sciences, Rockville, Maryland; [2]Armed Forces Institute of Pathology, Department of Genitourinary Pathology and [3]Urology Service, Department of Surgery, Walter Reed Army Medical Center, Washington, District of Columbia; and [4]U.S. Military Cancer Institute, Uniformed Services University of the Health Sciences, Bethesda, Maryland
Received 2/26/08; revised 4/25/08; accepted 4/28/08.

Grant support: Center for Prostate Disease Research and NIH grants RO1 DK065977 (S. Srivastava and G. Petrovics) and RO1 CA106653 (S. Srivastava and A. Dobi).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (http://clincancerres.aacrjournals.org/).

Y. Hu and A. Dobi contributed equally to this work.

The views expressed in this article are those of the authors and do not reflect the official policy of the Department of the Army, Department of Defense, or the U.S. Government.

Requests for reprints: Shiv Srivastava or Gyorgy Petrovics, Center for Prostate Disease Research, Department of Surgery, Uniformed Services University of the Health Sciences, Rockville, MD 20852. E-mail: ssrivastava@cpdr.org or gpetrovics@cpdr.org.

©2008 American Association for Cancer Research.

doi:10.1158/1078-0432.CCR-08-0531

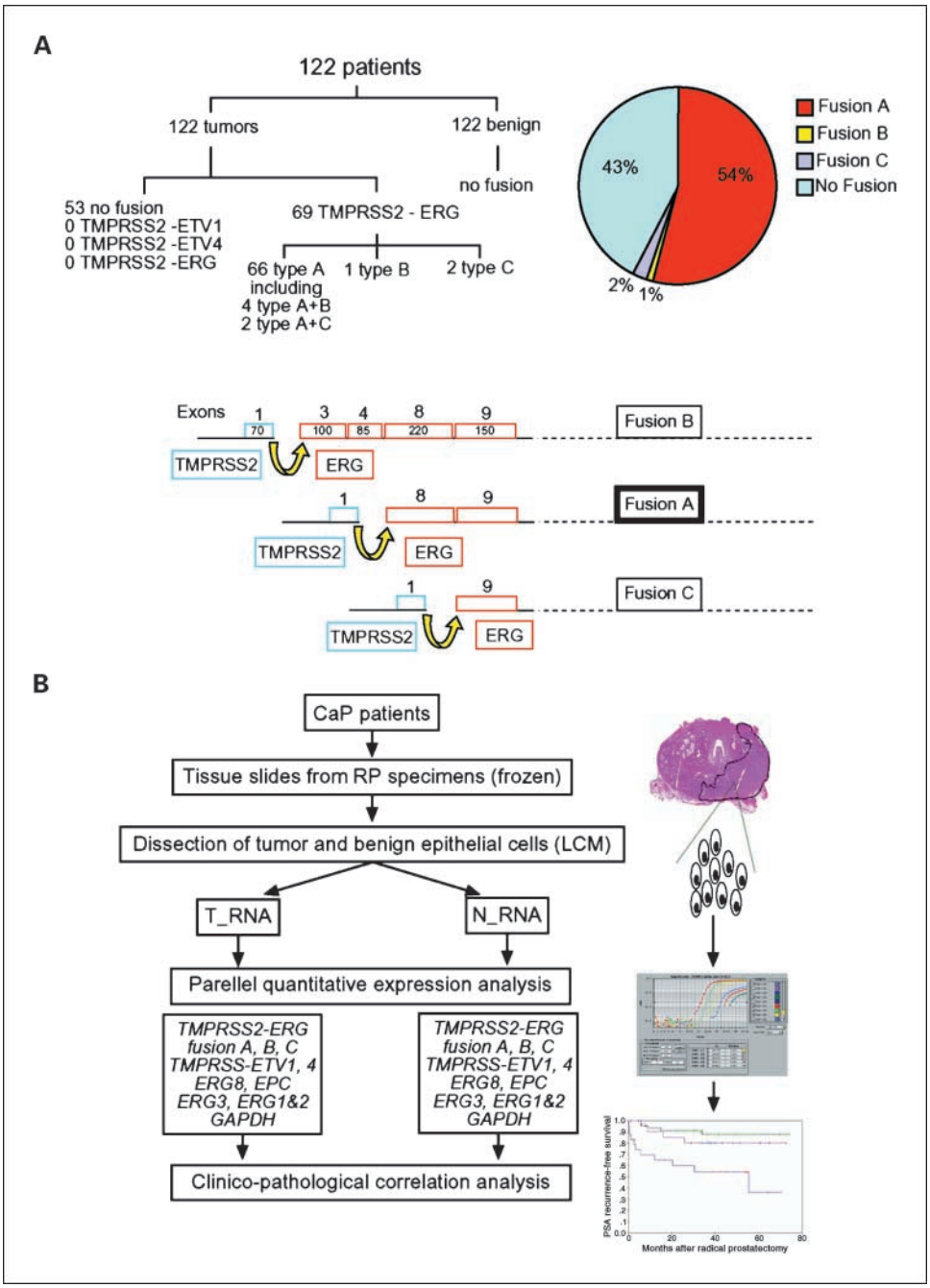


Fig. 1. Expression of *TMPRSS2-ERG* fusions in prostate cancer specimens. *A,* evaluation of the prevalence of *TMPRSS2-ERG* fusion A transcript junctions in prostate cancer patients (N = 122): Tumor and matching benign cells were assayed for *TMPRSS2-ERG, TMPRSS2-ETV1*, and *TMPRSS2-ETV4* fusion transcript junctions by quantitative reverse transcription-PCR. The schematic representation depicts the distribution of the various fusions in the patient specimens. The pie chart summarizes the *TMPRSS2-ERG* fusion junction types in the 122 tumor specimens. inset, *TMPRSS2-ERG* fusion A (*T1-E8*), B (*T1-E3*), and C (*T1-E9*) detected in this study (exon numbering is according to ref. 9). *B,* schematic representation of the experimental strategy and workflow of quantitative gene expression analysis in prostate tumor specimens.

In this study, we have cloned and sequenced full-length cDNAs from *TMPRSS2-ERG* fusion–positive prostate tumors, and from the VCaP cell line. We have identified two types of *TMPRSS2-ERG* cDNAs, one (type I) encoding full-length prototypical ERG protein (*ERG1*, *ERG2*, *ERG3*) and the other (type II) encoding a shorter version lacking the ETS domain (*ERG8* and a new variant, *TEPC*). We have further quantified and validated the expression of these *ERG* splice forms in a large cohort of prostate cancer specimens. The *ERG* exons at the *TMPRSS2-ERG* fusion junction have been the subject of a number of studies (fusion junction variants; refs. 5–7, 12–19). However, these exons are present in all *ERG* splice forms and do not identify specific splice variants. In recent *in vivo* models assessing the role of ERG in prostate cancer, only a type I splice variant, specifically NH_2-terminally truncated *ERG3*, was tested (20, 21). Intriguingly, the data presented here shows a more abundant expression of type II splice variants in prostate cancer cells. Our new findings on *ERG* splice variants in prostate cancer have promise in improving the understanding of *ERG* functions and its therapeutic targeting in prostate cancer, as well as in enhancing the detection of *ERG* alterations in clinical specimens.

Materials and Methods

Tissue specimens, laser capture microdissection, and quantitative gene expression analysis. The prostate tissue specimens used in this study

were obtained from radical prostatectomy procedures under an Institutional Review Board–approved protocol at Walter Reed Army Medical Center. Laser capture microdissection (LCM) of tumor and benign epithelial cells from optimum cutting temperature–embedded frozen tissues obtained from the radical prostatectomy specimens, RNA isolation from the LCM samples, and real-time quantitative reverse transcription-PCR (TaqMan) were essentially done as described previously (4, 22). The differentiation status of microdissected cells was recorded independently from the overall pathologic Gleason grade of the prostate, which was determined from whole-mounted, formalin-fixed, paraffin-embedded prostate specimens of each patient. The small amounts of tissue specimens (~ 5 mm^3) were obtained for optimum cutting temperature embedding from radical prostatectomy specimen before whole-mount prostate processing. Selection of specimens for LCM was primarily driven by the presence of sufficient amount of tumor cells for the LCM. The predominant tumor cell type (by differentiation) present in a frozen section was microdissected. Most of the time, but not always, the predominant differentiation grade in frozen tissue section represented the prevalent differentiation grade of the tumor cells in the prostate. Overall, 88.4% of LCM samples were collected from the primary Gleason pattern of the index tumor. Because of this, we have compared the quantitative gene expression of *TMPRSS2-ERG* splice variants to the differentiation grade of the microdissected cells, as well as to the overall differentiation grade of the of tumor cells in the prostate. Overall conclusions were similar by two-way comparisons but an increased statistically significant relationship was noted when the gene expression ratios of ERG I/II in LCM-RNAs was correlated with the differentiation grade of the LCM-dissected cells. TaqMan primers and probes are listed in the Supplementary data.

Detection of the *TMPRSS2-ERG* and *TMPRSS2-ETV* fusion transcripts was done essentially as described (5). The different *TMPRSS2-ERG* fusion junction types (A, B, and C) are described in Fig. 1A in a schematic diagram. All three fusion types have been previously described (5, 19). The expression of *GAPDH* was simultaneously analyzed as endogenous control, and the target gene expression in each sample (in duplicates) was normalized to *GAPDH*. RNA samples without reverse transcription were included as the negative control in each assay.

Generation and screening of cDNA library from prostate tumors. For the generation of the cDNA library, frozen tumor tissues from index tumors of six patients were selected based on available tissue size (over 30 mg), highest tumor cell content (over 70%), and the presence of *TMPRSS2-ERG* fusion transcripts by reverse transcription-PCR. Polyadenylated RNA was isolated from the optimum cutting temperature–embedded frozen tumor tissues. A cDNA library was generated from the pooled RNA (Lofstrand Laboratories) and cloned into the *Xho*I-*Eco*RI sites of lambdaZAP Express vector (Stratagene). Screening of the expression library was carried out according to the protocol described by the manufacturer (Stratagene). The primary library of about 400,000 plaques were screened by *ERG2* probe (NM_004449; cDNA obtained from Dr. Dennis Watson, Medical University of South Carolina, Charleston, SC) and found 84 hybridized with different intensities. The positive plaques were further screened for the presence of *TMPRSS2* fusions by fusion-specific PCR (5). A total of 12 plaques showed

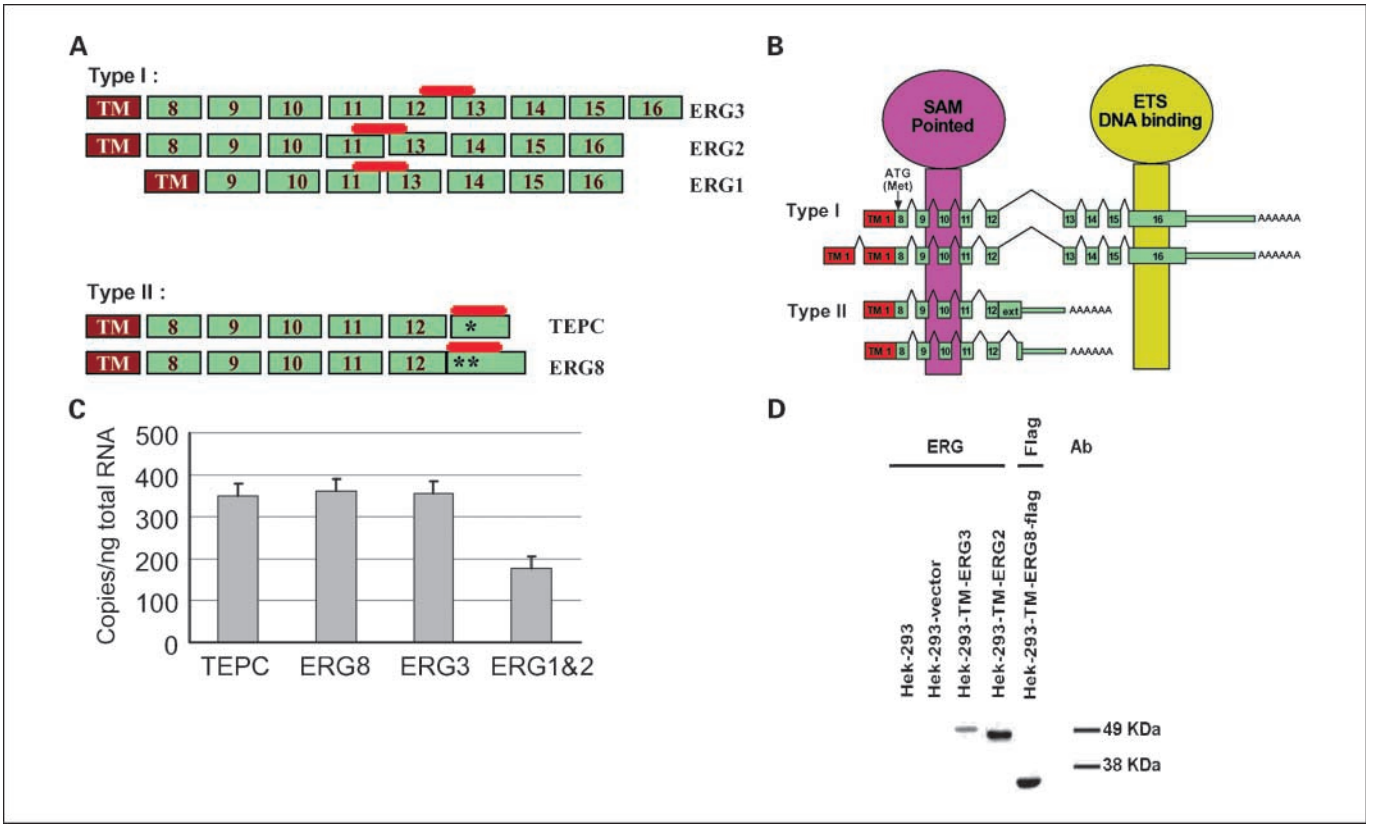


Fig. 2. *ERG* splice forms in prostate tumors and their expression in the VCaP prostate cancer cell line. *A,* schematic representation of full-length type I and type II *ERG* transcripts expressed in prostate cancer cells with *TMPRSS2-ERG* fusion. Numbered boxes, *ERG* exons (9); boxes with * and **, unique regions of *TEPC* and *ERG8*, respectively. Solid lines above the exons, TaqMan primers and probes used for the detection of the *ERG* splice variants. *B,* type I transcripts code for both transactivation (*SAM Pointed*) and DNA-binding (*ETS*) domains. In contrast, type II variants lack the coding sequence for the DNA-binding domain. The relative positions coding for the two major functional domains of ERG protein are shown in type I and type II splice variants. *C,* columns, copy numbers of the *ERG* splice forms in VCaP cells determined by TaqMan quantitative reverse transcription-PCR. The median of three experiments using triplicates are shown. *D,* protein products expressed from *TMPRSS2-ERG2, TMPRSS2-ERG3*, and *TMPRSS2-ERG8* clones transiently transfected into HEK293 cells are shown by Western blot analysis. Anti-Flag antibody was used for the detection of Flag-tagged ERG8 protein.

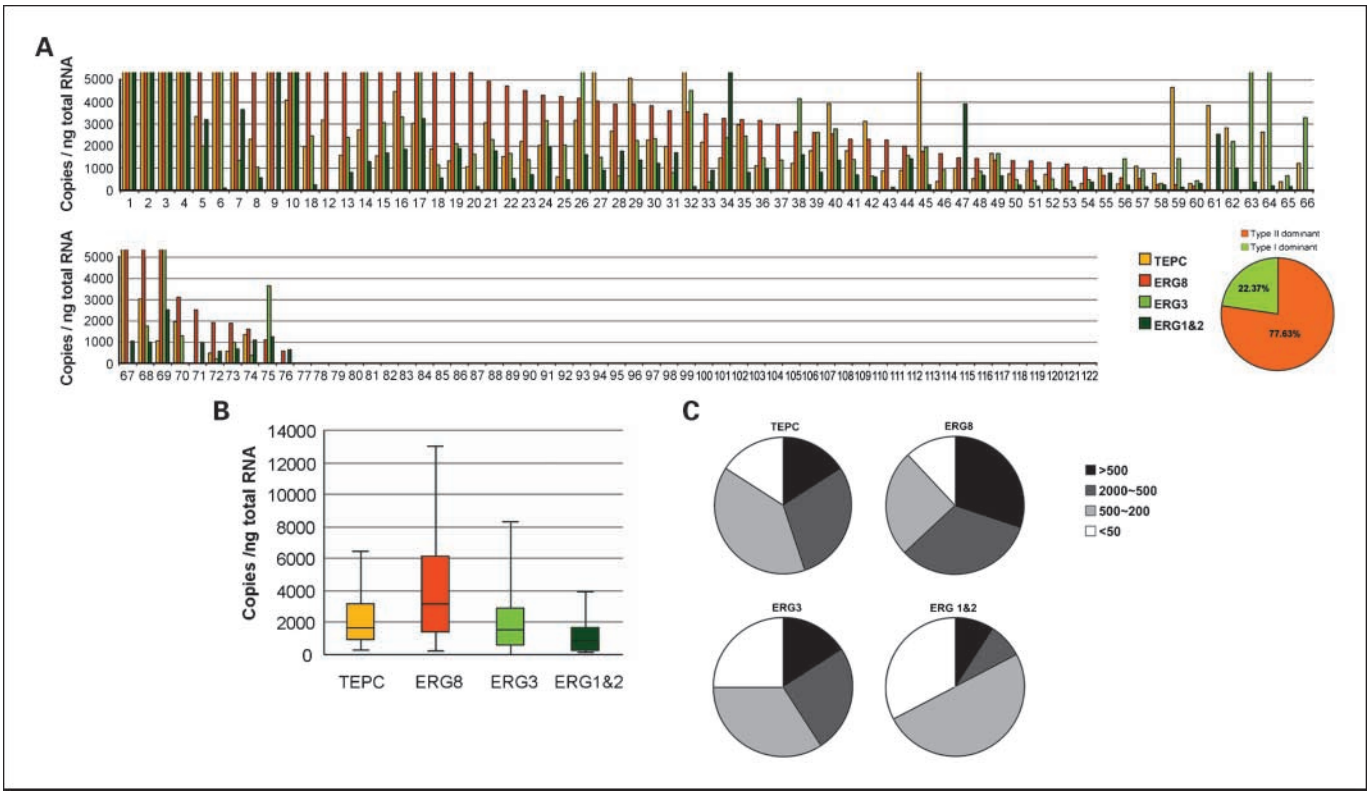


Fig. 3. Quantitative expression of *ERG* splice variants in prostate cancer patients. *A,* quantitative expression of *ERG* splice forms *ERG8, TEPC, ERG3*, and *ERG1&2* (*columns with different colors*) were determined in microdissected tumor cells of prostate cancer patients ($N = 122$). The graph depicts relative expression levels (copy number/ng total RNA, normalized to *GAPDH*) in patients with (*top*, $n = 66$) and with no detectable (*bottom*, $n = 56$) *TMPRSS2-ERG* fusion A transcript junction. Due to wide dynamic range of copy numbers, the values above 5,000 are not shown here (the range of expression for each splice variant is depicted in *B*). In the pie chart, the percentage of prostate cancer patients are represented either with higher (~77%) or with lower (~23%) expression level of type II than type I *ERG* transcripts. *B,* relative abundance of various *ERG* splice forms (*ERG8, TEPC, ERG3*, and *ERG1&2*) is depicted by box plots representing the copy numbers determined in 76 prostate cancer patients overexpressing *ERG*. *C,* pie charts illustrate the distribution of prostate cancer patients with various expression levels (copy number/ng total RNA) of *ERG8, TEPC, ERG3*, and *ERG1&2* splice forms in prostate cancer cells.

amplification. Detailed cDNA sequence analysis revealed the presence of two types of *TMPRSS2-ERG* fusion transcripts. Within the positively identified plaques, three represented type I (with both SAM domain and DNA-binding ETS domain) and five type II (without ETS domain). Fusion-positive type I– and type II–containing phages were amplified with T3 and T7 primers, subcloned into TOPO vector (Invitrogen), and verified by DNA sequencing.

Cell culture and Western blot. The prostate cancer cell line VCaP, which has type A *TMPRSS2-ERG* fusion (5), and human embryonic kidney HEK 293 cells were obtained from the American Type Culture Collection. Cells were cultured according to the provider's instructions[5] and harvested upon confluence of 70%. RNA was isolated by RNAzol B method (Tel-Test, Inc.). *TMPRSS2-ERG2*, *TMPRSS2-ERG3*, and *TMPRSS2-ERG8* (Flag-tagged) constructs were cloned from prostate cancer cDNA library into pIRES-EGFP plasmid vector (Clontech) and were verified by DNA sequencing. HEK293 cells transfected with the constructs were lysed in M-PER mammalian protein extraction reagent (Pierce) supplemented with protease and phosphatase inhibitor cocktails (Sigma). ERG2 and ERG3 proteins were detected by Western blot (NuPAGE Bis-Tris gel, Invitrogen) using immunoaffinity-purified anti-ERG peptide polyclonal antibody prepared in our laboratory (DFHGI AQALQ PHPPE SSLYK YPSDL PYMGS YHAHP QKMNF VAPHP PAL). The tagged ERG8 protein was detected by Flag-tag antibody (Sigma).

Statistical analyses of clinical and gene expression data. Measures of central tendency (median) and dispersion (range) are used to describe continuously measured patient characteristics, whereas frequencies and percentages are used to describe categorical patient characteristics. χ^2 and Fisher's exact tests were conducted to compare *TMPRSS2-ERG* splice variant transcript expression across patient clinical and demographic characteristics. P values <0.05 are considered statistically significant.

Results

Quantitative analysis of TMPRSS2-ERG expression in prostate tumors. Quantitative analyses of the transcript levels of various *TMPRSS2-ETS* fusion genes were done in LCM matched benign and tumor epithelium of prostate cancer specimens (122 patients; 244 specimens; Fig. 1A). The demographic, clinical, and pathologic variables of the patient cohort are summarized in Supplementary Table S1. The workflow of LCM and quantitative reverse transcription-PCR analysis is summarized in Fig. 1B. The most frequently observed *TMPRSS2-ERG* fusion transcript junctions (6, 7) were detected in 57% of the patients, and among these 95% expressed *TMPRSS2-ERG* fusion type A (Fig. 1A). Fusions with other ETS family members, such as *TMPRSS2-ETV1* or *TMPRSS2-ETV4*, were not detected in this cohort. No fusions were detected in matched benign prostate epithelial cells dissected from the same prostate.

Identification of full-length TMPRSS2-ERG transcripts in prostate tumors. To investigate the nature of *TMPRSS2-ERG*–encoded proteins in prostate cancer, a cDNA library was generated from RNA pooled from six prostate tumors with

[5] http://www.atcc.org

TMPRSS2-ERG fusion. Screening of the library (see flow chart in Supplementary Fig. S1) by both *ERG* and *TMPRSS2* probes resulted in the identification of the following *ERG* splice variants: *ERG1* (M21535), *ERG2* (NM004449), *ERG3* (NM182918), *ERG8* (AY204742), and *TEPC*, a novel splice variant (EU432099; Fig. 2A). *ERG1*, *ERG2*, and *ERG3* contain both SAM (pointed) and ETS (DNA-binding) domain (type I); however, *ERG8* and *TEPC* lack the ETS domain (type II; Fig. 2B). Among the positively identified cDNA library clones, 30% were type I and 70% were type II. Both types of *ERG* transcripts are expressed in VCaP cells, a human prostate cancer cell line derived from vertebral metastasis that harbors *TMPRSS2-ERG* fusion, with the type II transcripts being more abundant (Fig. 2C).

HEK293 cells were transiently transfected with the *TMPRSS2-ERG2*, *TMPRSS2-ERG3*, and *TMPRSS2-ERG8* constructs and the expressed ERG proteins were detected by Western blot showing the expected molecular weight of type I and type II proteins (Fig. 2D). For the detection of type I splice forms (ERG2 and ERG3), an anti-peptide polyclonal ERG antibody was used, which was developed in our laboratory. ERG8, a type II splice form, was Flag-tagged and detected by anti-Flag antibody, because ERG8 lacks part of our ERG peptide epitope.

Relative abundance of type II ERG splice forms in tumor cells of prostate cancer patients. Quantitative expression of the *ERG* splice variants were determined in microdissected tumor cells of 122 prostate cancer patients: 66 with *TMPRSS2-ERG* fusion A transcript junction and 56 with no detectable fusion A transcript (Fig. 3A). At least two or more *ERG* splice variants were detectable in all *TMPRSS2-ERG* fusion A–positive prostate cancer patients. *ERG8* and *TEPC* represented the most abundant *ERG* splice forms analyzed (Fig. 3B) and were detected in 65 of 66 *TMPRSS2-ERG* fusion A expression–positive patients (Fig. 3A). Expression of at least two of the *ERG* splice forms was detected in 10 of 56 fusion A expression–negative cases. Three of these tumors were positive for *TMPRSS2-ERG* fusion types B or C. It is likely that other such tumors may harbor other *TMPRSS2-ERG* fusion junctions. Thus, quantitative analysis of *ERG* splice variants, especially *ERG8* and *TEPC*, provide a reliable surrogate for *TMPRSS2-ERG* fusion in prostate cancer, and in addition it detects *ERG* overexpression even if the fusion junction type is unknown. The order of median abundance (copies/ng total RNA) of *ERG* splice forms in prostate cancer cells of 76 patients with detectable *ERG* expression was *ERG8* (~3,200) > *TEPC* (~1,800) > *ERG3* (~1,500)> *ERG1&2* (~800; Fig. 3B). Overall, the type II splice variants (with no ETS domain) were present in higher copy numbers in prostate cancer cells than the type I splice forms (Fig. 3B and C), and 77% of *ERG*-positive prostate cancer patients tested have more copies of type II than type I splice forms (Fig. 3A). We conclude that quantitative detection of *ERG* splice variants, especially *ERG8* and *TEPC*, may provide increased sensitivity in assessing overall frequency of *TMPRSS2-ERG* fusions in prostate cancer cells.

Expression of ERG splice forms in relation to clinicopathologic variables of prostate cancer patients. In comparison with

Table 1. Correlation of *ERG* splice variant expression and type I/type II ratio with clinicopathologic characteristics

Clinicopathologic characteristics	*ERG* splice variant expression			Type I/ type II ratio		
	No (46)	Yes (76)	*P**	*n* (76)	Median	*P* †
Race			0.0345			0.1142
Caucasian	29 (33%)	59 (67%)		59	0.51	
African American	15 (56%)	12 (44%)		12	0.33	
Family history			0.6312			0.3259
No	28 (38%)	45 (62%)		45	0.52	
Yes	10 (33%)	20 (67%)		20	0.51	
Pathologic T stage			0.5318			0.2541
pT_2	14 (36%)	25 (64%)		25	0.44	
pT_3	29 (42%)	40 (58%)		40	0.51	
Pathologic Gleason sum			0.0023			0.0323
2-6	9 (24%)	29 (76%)		29	0.35	
7	20 (36%)	36 (64%)		36	0.46	
8-10	14 (70%)	6 (30%)		6	0.70	
LCM differentiation			0.0058			0.0067
Well	31 (32%)	66 (68%)		65	0.45	
Poorly	15 (62%)	9 (38%)		9	0.76	
Margin status			0.9436			0.0032
Negative	28 (38%)	45 (62%)		45	0.39	
Positive	16 (39%)	25 (61%)		25	0.57	
PSA recurrence			0.7312			0.0456
No	33 (37%)	56 (63%)		56	0.42	
Yes	11 (41%)	16 (59%)		16	0.61	

NOTE: *ERG* splice variant expression (N = 122): In comparison with patients with no detectable expression of *ERG* in their prostate cancer cells (n = 46), the *ERG*-positive patient cohort (n = 76) has a decreased proportion of patients with high Gleason grade, poor prostate cancer cell differentiation, and African American ethnicity. Type I/type II ratio (n = 76): The ratio of *ERG* type I/type II splice variants in prostate cancer cells is increased in patients with poor tumor cell differentiation and with prostate-specific antigen recurrence.
Abbreviation: PSA, prostate-specific antigen.
*Two-sided test, $P < 0.05$ was considered statistically significant.
† One-sided test, $P < 0.05$ was considered statistically significant.

prostate cancer patients with no detectable *ERG* expression ($n = 46$), the *ERG* expression–positive patient cohort ($n = 76$) has a smaller proportion of patients with high pathologic Gleason grade (8–10), poor prostate cancer cell differentiation, or African American ethnicity (Table 1). The levels of type I or type II *ERG* splice forms in the cohort of prostate cancer patients with *ERG* expression ($n = 76$) did not show significant correlations with clinicopathologic variables. However, there was a trend of correlation of higher copy number ratio of type I over type II splice forms with poor differentiation of prostate cancer cells, higher pathologic Gleason sum, positive margin, and biochemical recurrence (Table 1).

Discussion

ERG overexpression as a result of *TMPRSS2-ERG* fusion represents a highly prevalent oncogenic alteration in prostate cancer. Remarkable progress has been made in just over 2 years in establishing the diagnostic and prognostic features of *TMPRSS2-ERG* fusion in prostate cancer (6, 7). Despite the large body of data on the *TMPRSS2-ERG* fusion junctions, virtually nothing is known about the full-length *TMPRSS2-ERG* transcripts, including the existence and relative abundance of specific splice variants in human prostate tumors. However, splice variants of numerous genes, e.g., *androgen receptor*, *fibroblast growth factor receptor*, *survivin*, and *MDM2*, are known to play critical roles in various human cancers (10, 11).

This study establishes the nature of full-length *TMPRSS2-ERG* transcripts and encoded proteins in prostate cancer cells. In addition to expected full-length *TMPRSS2-ERG* transcripts, we have identified relatively abundant *ERG* splice forms with unique 3′ sequences that lack a conserved region coding for the DNA binding ETS domain. Parallel quantitative analyses of *ERG* splice variants in precisely microdissected cells from well-defined histologic features of the tumor provided accurate data with respect to the presence, abundance, and distribution of various *ERG* splice forms in prostate cancer in relation with clinicopathologic status.

Monitoring the expression of *ERG* splice variants, we detected more prostate cancer cases than by monitoring the fusion transcript junctions, likely because unknown or undetected fusions are present in a subset of cases. Furthermore, the number of various fusion junctions in prostate cancer is far more than the number of *ERG* splice variants.

Recent reports revealed that specific junction types of *TMPRSS2-ERG* fusion transcripts, genomic deletions, or the presence of *TMPRSS2-ERG* fusion are associated with poor prognosis (reviewed in refs. 6, 7). However, others reported that fusion-positive tumors were associated with lower Gleason grade and/or better disease outcome (12). In this study, we found that compared with patients with no detectable expression of *ERG* in their prostate cancer cells, the *ERG* expression–positive patient cohort has a decreased proportion of patients with high Gleason grade, poor prostate cancer cell differentiation, and African American ethnicity ($N = 122$). This is in agreement with our previous study on *ERG* expression in prostate cancer (4). Lower or no *ERG* expression in a subset of aggressive tumors with *TMPRSS2-ERG* fusion may reflect attenuation of androgen signaling pathway during prostate cancer progression (23). The levels of type I or type II *ERG* splice forms did not show significant correlations with clinicopathologic variables. It will be useful to combine multiple approaches, including quantitative assessment of *TMPRSS2-ERG* expression levels, evaluation of genomic rearrangements, and different types of transcripts in multicenter cohort to confirm prognostic values of qualitative and quantitative aspects of *ERG* alterations in prostate cancer.

The diversity of *TMPRSS2-ERG* fusion transcripts has recently been emphasized focusing on the fusion junction region of the transcripts (13, 14) and by using exon arrays (15) that did not allow for the discovery of the type II splice variants described here. Our results highlight the importance of understanding the expression and distribution of full-length splice forms of *ERG*, including variants with no DNA binding domain, in the tumor cells. Our data show a trend of correlation of relatively more type I over type II splice forms, with less favorable pathology and outcome that need to be confirmed in a larger patient cohort. The heterogeneity of *TMPRSS2-ERG* rearrangements in multifocal prostate cancer reported by our group and others (18, 19) further adds to the complexity of understanding the roles of *ERG* in prostate cancer.

In conclusion, this study establishes two major types of full-length transcripts from the *TMPRSS2-ERG* locus in prostate cancer. Further, we establish the protein products translated from type I and type II transcripts. The presence of these specific *ERG* splice forms, especially the more abundant type II splice forms, may provide new opportunities in as prostate cancer biomarker. Finally, overall status of the type I and II forms in prostate cancer cells, such as the ratio of their expression levels, has potential to enhance our understanding of the biology of prostate tumors with *TMPRSS2-ERG* fusion.

Disclosure of Potential Conflicts of Interest

None of the authors have competing financial interests.

Acknowledgments

We thank Drs. Vasantha Srikantan and Govindan Vaidyanathan for their contributions to this study; Drs. Dennis K. Watson (Medical University of South Carolina) and Charles J. Bieberich (University of Maryland, Baltimore Campus) for their valuable advice and comments during the progress of this study; and Amina Ali and Yinghui Shi and for their superb technical support.

References

1. De Marzo AM, Platz EA, Sutcliffe S, et al. Inflammation in prostate carcinogenesis. Nat Rev Cancer 2007; 7:256–69.
2. Richter E, Srivastava S, Dobi A. Androgen receptor and prostate cancer. Prostate Cancer Prostatic Dis 2007;10:114–8.
3. Kallioniemi O. Functional genomics and transcriptomics of prostate cancer: promises and limitations. BJU Int 2005;96 Suppl 2:10–5.
4. Petrovics G, Liu A, Shaheduzzaman S, et al. Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome. Oncogene 2005;24: 3847–52.
5. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644–8.
6. Turner DPWatson DK. ETS transcription factors: oncogenes and tumor suppressor genes as therapeutic targets for prostate cancer. Expert Rev Anticancer Ther 2008;8:33–42.
7. Demichelis F, Rubin MA. TMPRSS2-ETS fusion prostate cancer: biological and clinical implications. J Clin Pathol 2007;60:1185–6.
8. Seth A, Watson DK. ETS transcription factors and their emerging roles in human cancer. Eur J Cancer 2005;41:2462–78.
9. Owczarek CM, Portbury KJ, Hardy MP, et al. Detailed

mapping of the ERG-ETS2 interval of human chromosome 21 and comparison with the region of conserved synteny on mouse chromosome 16. Gene 2004;324:65–77.

10. Venables JP. Unbalanced alternative splicing and its significance in cancer. Bioessays 2006;28:378–86.

11. Steinman HA, Burstein E, Lengner C, et al. An alternative splice form of Mdm2 induces p53-independent cell growth and tumorigenesis. J Biol Chem 2004;279:4877–86.

12. Tu JJ, Rohan S, Kao J, et al. Gene fusions between TMPRSS2 and ETS family genes in prostate cancer: frequency and transcript variant analysis by RT-PCR and FISH on paraffin-embedded tissues. Mod Pathol 2007;20:921–28.

13. Clark J, Merson S, Jhavar S, et al. Diversity of TMPRSS2-ERG fusion transcripts in the human prostate. Oncogene 2007;26:2667–73.

14. Liu W, Ewing CM, Chang BL, et al. Multiple genomic alterations on 21q22 predict various TMPRSS2/ERG fusion transcripts in human prostate cancers. Genes Chromosomes Cancer 2007;46:972–80.

15. Jhavar S, Reid A, Clark J, et al. Detection of TMPRSS2-ERG translocations in human prostate cancer by expression profiling using GeneChip Human Exon 1.0 ST arrays. J Mol Diagn 2008;10:50–7.

16. Rajput AB, Miller MA, De Luca A, et al. Frequency of the TMPRSS2:ERG gene fusion is increased in moderate to poorly differentiated prostate cancers. J Clin Pathol 2007;60:1238–43.

17. Wang J, Cai Y, Ren C, Ittmann M. Expression of variant TMPRSS2/ERG fusion messenger RNAs is associated with aggressive prostate cancer. Cancer Res 2006;66:8347–51.

18. Mehra R, Han B, Tomlins SA, et al. Heterogeneity of TMPRSS2 gene rearrangements in multifocal prostate adenocarcinoma: molecular evidence for an independent group of diseases. Cancer Res 2007;67:7991–5.

19. Furusato B, Gao CL, Ravindranath L, et al. Mapping of TMPRSS2-ERG fusions in the context of multi-focal prostate cancer. Mod Pathol 2008;21:67–75.

20. Klezovitch O, Risk M, Coleman I, et al. A causal role for ERG in neoplastic transformation of prostate epithelium. Proc Natl Acad Sci U S A 2008;105:2105–10.

21. Tomlins SA, Laxman B, Varambally S, et al. Role of the TMPRSS2-ERG gene fusion in prostate cancer. Neoplasia 2008;10:177–88.

22. Petrovics G, Zhang W, Makarem M, et al. Elevated expression of PCGEM1, a prostate-specific gene with cell growth-promoting function, is associated with high-risk prostate cancer patients. Oncogene 2004;23:605–11.

23. Sterbis JR, Gao C, Furusato B, et al. Higher expression of the androgen-regulated gene PSA/HK3 mRNA in prostate cancer tissues predicts biochemical recurrence-free survival. Clin Cancer Res 2008;14:758–63.

KIMBROUGH UROLOGICAL SEMINAR-WASHINGTON DC, JAN 11-16, 2009.

Splice Variants and Functions of TMPRSS2-ERG Fusion, A Common Genomic Alteration in Prostate Cancer.

Albert Dobi, Chen Sun, Ying Hu, Rajesh Thangapazham, Ahmed Mohamed, CPT Eric J. Whitman, CPT Dorotha Hawksworth, Bungo Furusato, Shyh-Han Tan, Atekelt Y. Tadase, Lakshmi Ravindranath, Jennifer Cullen, Yongmei Chen, Gyorgy Petrovics, Taduru Sreenath, Isabell A. Sestrehenn, COL David G. McLeod, Shiv Srivastava. Urology Service and Center for Prostate Disease Research (CPDR), Walter Reed Army Medical Center, Washington, DC; Center for Prostate Disease Research, Department of Surgery, Uniformed Services University of the Health Sciences, Rockville, MD; Armed Forces Institute of Pathology, Washington, DC.

Won Best Poster Award

Introduction and Objectives *ERG* oncogene is overexpressed in the majority of prostate cancers due to genomic rearrangements placing the ERG gene under the control of the androgen regulated TMPRSS2 promoter. However, the biological functions of *ERG* in CaP are not well understood. This study was designed to evaluate the function of ERG and to reveal the forms of ERG in human prostate tumors.

Materials and Methods: ERG was knocked down in TMPRSS2-ERG expressing VCaP cells. Cell proliferation and tumor growth was assayed by in vitro and in vivo models, respectively. Gene expression changes were evaluated by microarray and by quantitative PCR. ERG splice variants were cloned from a Lambda-Zap cDNA library from TMPRSS2-ERG expressing prostate tumors from six patients.

Results: We demonstrated that ERG activates C-MYC oncogene and negatively regulates the expression of prostate differentiation markers in prostate cancer. The predominant form of ERG in human prostate tumors is a novel type (type II) form that lacks DNA binding domain.

Conclusions: The primary consequence of ERG overexpression in prostate tumors is the suppression of prostate differentiation genes. ERG positively regulates C-MYC oncogene contributing to the suppression of differentiated cellular phenotype. Detection of prostate cancer by Type I and II variants in human prostate tissues is superior to the detection of tumors by the TMPRSS2-ERG fusion junction.

SOCIETY FOR BASIC UROLOGY RESEARCH, PHOENIX, AZ, NOV 20-23, 2008.

TMPRSS2-ERG Splice Variants in Prostate Cancer

Shyh-Han Tan[1], Ying Hu[1], Albert Dobi[1], Taduru Sreenath[1], Christopher Cook[1], Atekelt Y. Tadese[1], Lakshmi Ravindranath[1], Jennifer Cullen[1], Bungo Furusato[1,2], Yongmei Chen[1], Rajesh Thangapazham[1], Ahmed Mohamed[1], Chen Sun[1], Isabell A. Sesterhenn[2], David G. McLeod[1,3,4], Gyorgy Petrovics[1] and Shiv Srivastava[1,4]. [1]Center for Prostate Disease

Research, Dept.of Surgery, Uniformed Services University of the Health Sciences (USUHS), Rockville MD 20852, [2] Armed Forces Institute of Pathology, Dept. of Genitourinary Pathology, Washington DC 20307, [3] Urology Service, Dept.of Surgery, Walter Reed Army Medical Center, Washington DC 20307, [4] US Military Cancer Institute, USUHS, Bethesda, MD 20814

Introduction and Objective: Recent studies have confirmed that high proportion (>70%) of prostate cancers (CaP) harbor over-expression of the ETS related genes (ERG, ETV1 or ETV4) as a result of the fusion of an androgen receptor regulated promoter element of the TMPRSS2 gene to the protein coding sequence of ETS related genes. TMPRSS2-ERG fusion junctions have been extensively studied in CaP. However, the nature of full length transcripts and encoded proteins remains unknown. This study focuses on qualitative and quantitative features of full length TMPRSS2-ERG transcripts in CaP.

Methods: Full length TMPRSS2-ERG transcripts were cloned and sequenced from a cDNA library generated from pooled RNA of six TMPRSS2-ERG fusion positive prostate tumors. The encoded ERG proteins were analyzed in HEK293 cells. Copy numbers of TMPRSS2-ERG splice variants were determined by QRT-PCR in laser capture micro-dissected CaP cells.

Results: Two types of TMPRSS2-ERG cDNAs were identified: type I, that encodes full length prototypical ERG protein (ERG1, ERG2, ERG3) and type II, encoding truncated ERG proteins lacking the ETS domain (ERG8 and a new variant, TEPC). Among ERG negative patients, 4 (8.7%) of them were found ETV1 over-expression. In microdissected prostate tumor cells from 122 patients, relative abundance of these variants was in the following order: ERG8 > TEPC > ERG 3 > ERG1/2 with combined over-expression rate of 62.3% in CaP. Increased ratio of type I over type II splice forms showed a trend of correlation with less favorable pathology and outcome. The proportion of protein expression of ERG8 in relation to the other ERG isoforms remains to be analyzed.

Conclusions: Qualitative and quantitative features of specific ERG splice variants defined here promise to enhance the utility of ERG as a biomarker and therapeutic target in CaP.

ADVANCES IN PROSTATE CANCER (AACR) SAN DIEGO, CA JAN 21-24, 2009.

Expression and Sub-cellular Localization of Predominant TMPRSS2-ERG Transcripts with and without ETS domain in Prostate Cancer

Taduru L. Sreenath, Ying Hu, Albert Dobi, Atekelt Tadese, Shyh-Han Tan, Lakshmi Ravindranath, Bungo Furusato, Youngmei Chen, Rajesh Thangapazham, Amina Ali, Ahmed Mohamed, Chen Sun, Isabell Sesterhenn, David G. McLeod, Gyorgy Petrovics, Shiv Srivastava.

Introduction: High prevalence of ETS Related Gene (ERG) over expression in prostate cancer cells due to TMPRSS2-ERG fusions suggests a causal role for ERG protein(s) in the neoplastic process. ERG is classified as a member of ETS transcription factor gene family, consisting of a protein-protein interaction domain (SAM), and a DNA binding domain (ETS). Several types of fusions between TMPRESS2 and ERG have been identified and few of these fusions result into deletion of ~33-39 amino acids at the N-terminal end of ERG protein(s). Although, TMPRSS2-ERG fusion junctions at the 5'-end to ERG have been extensively studied in prostate cancer, the presence and the nature of full-length transcripts, their encoded protein products, their sub-cellular localization and their relevance to the clinicopathological stages is not well understood. This study focuses on qualitative and quantitative features of full-length TMPRSS2-ERG transcripts in prostate cancer and the sub-cellular localization of the encoded proteins.

Experimental procedures: cDNA library was generated from pooled RNA of six TMPRSS2-ERG fusion positive prostate tumors. Cloning and sequencing of the identified transcripts were performed using standard molecular biology techniques. Full-length clones identified were expressed in HEK293 cells to further characterize the proteins and their sub-cellular localization. Differential expression of TMPRSS2- ERG splice variants were determined by quantitative reverse transcription-PCR in laser capture micro dissected prostate cancer cells to correlate their ratios with clinicopathological stages.

Results: We identified two types of TMPRSS2-ERG cDNAs by prostate cancer library screening using human TMPRSS2 and ERG cDNA fragments as probes. Type I, which encodes full-length ERG protein consisting SAM and ETS domains (ERG1, ERG2, ERG3), and type II, encoding ERG proteins lacking the ETS domain (ERG8 and a new variant,TEPC). Quantitative RT-PCR analyses of fusion transcripts from LCM dissected prostate tumor cells of 122 patients have shown the presence of various transcripts. The relative abundance of these variants was in the following order: ERG8 > TEPC > ERG 3 > ERG1/2 with a combined over expression rate of 62.3% in prostate cancer. Increased ratio of type I over type II splice forms showed a trend of correlation with poorly differentiated pathology /high Gleason score and outcome. Isolated type I and type cDNAs were expressed in HEK293 cells to study the sub-cellular localization by immunocytochemistry. Our studies revealed the presence of TMPRSS2-ERG3 (type I) in the nucleus and TMPRSS2-ERG8 (type II) in the cytoplasm.

Conclusions: The ratios of type I/type II ERG transcript expression by quantitative analysis correlated with poorly differentiated tumors. Qualitative and quantitative features of specific ERG transcripts defined here promise to enhance the utility of ERG as a biomarker and therapeutic target in prostate cancer.

AMERICAN ASSOCIATION FOR CANCER RESEARCH, DENVER CO, APRIL 18-22, 2009

Predominant TMPRSS2-ERG Transcripts in Prostate Cancer

Ying Hu, Albert Dobi, Taduru Sreenath, Atekelt Tadese, Shyh-Han Tan, Lakshmi Ravindranath, Bungo Furusato, Yongmei Chen, Rajesh Thangapazham, Amina Ali, Ahmed Mohamed, Chen Sun, Isabell Sesterhenn, David McLeod, Gyorgy Petrovics, Shiv Srivastava.

Introduction: High prevalence of ETS Related Gene (ERG) over expression in prostate cancer cells due to TMPRSS2-ERG fusions suggests a causal role for ERG protein(s) in the neoplastic process. ERG is classified as a member of ETS transcription factor gene family, consisting of a protein-protein interaction domain (SAM), and a DNA binding domain (ETS). Several types of fusions between TMPRSS2 and ERG have been identified and few of these fusions result into deletion of ~33-39 amino acids at the N-terminal end of ERG protein(s). Although, TMPRSS2-ERG fusion junctions at the 5'-end to ERG have been extensively studied in prostate cancer, the presence and the nature of full-length transcripts, their encoded proteins, their sub-cellular localization and their relevance to the clinicopathological stages is not well understood. This study focuses on qualitative and quantitative features of full-length TMPRSS2-ERG transcripts in prostate cancer and the sub-cellular localization of the encoded proteins.
Experimental procedures: cDNA library was generated from pooled RNA of six TMPRSS2-ERG fusion positive prostate tumors. Cloning and sequencing of the identified transcripts were performed using standard molecular biology techniques. Full-length clones identified were expressed in HEK293 cells to further characterize the proteins and their sub-cellular localization. Differential expression of TMPRSS2- ERG splice variants were determined by quantitative reverse transcription-PCR in laser capture micro dissected prostate cancer cells to correlate their ratios with clinicopathological stages.
Results: We identified two types of TMPRSS2-ERG cDNAs by prostate cancer library screening using human TMPRSS2 and ERG cDNA fragments as probes. Type I, which encodes full-length ERG protein consisting SAM and ETS domains (ERG1, ERG2, ERG3), and type II, encoding ERG proteins lacking the ETS domain (ERG8 and a new variant,TEPC). Quantitative RT-PCR analyses of fusion transcripts from LCM dissected prostate tumor cells of 122 patients have shown the presence of various transcripts. The relative abundance of these variants was in the following order: ERG8 > TEPC > ERG 3 > ERG1/2 with a combined over expression rate of 62.3% in prostate cancer. Increased ratio of type I over type II splice forms showed a trend of correlation with poorly differentiated pathology /high Gleason score and outcome. Isolated type I and type cDNAs were expressed in HEK293 cells to study the sub-cellular localization by immunocytochemistry. Our studies revealed the presence of TMPRSS2-ERG3 (type I) in the nucleus and TMPRSS2-ERG8 (type II) in the cytoplasm.
Conclusions: The ratios of type I/type II ERG transcript expression correlated with poorly differentiated tumors. The features of specific ERG transcripts defined here promise to enhance the utility of ERG as a biomarker and therapeutic target in prostate cancer.

UNIFORMED SERVICES UNIVERSITY OF THE HEALTH SCIENCES RESEARCH WEEK, MAY 11-13

Differential Expressions and Sub-cellular Localization of TMPRSS2-ERG Splice Variants in Prostate Cancer

Taduru Sreenath, Ying Hu, Albert Dobi, Atekelt Y Tadese, Shyh-Han Tan, Lakshmi Ravindranath, Bungo Furusato, Yongmei Chen, Rajesh Thangapazham, Amina Ali, Ahmed Mohamed, Chen Sun, Isabell Sesterhenn, David G McLeod, Gyorgy Petrovics, Shiv Srivastava, Center for Prostate Disease Research, 1530 East Jefferson Street, Rockville, Maryland

ETS Related Gene (*ERG*) shares sequence homology to the E26 avian erythroblastosis retrovirus and is classified as a member of *ETS* transcription factor gene family. High prevalence of ERG over expression in prostate cancer cells due to TMPRSS2-ERG fusions suggest for causal roles of ERG protein in the neoplastic process. Although, TMPRSS2-ERG fusion junctions have been extensively studied in prostate cancer, the presence and the nature of full-length transcripts, their encoded protein products, their sub-cellular localization and their relevance to the clinicopathological stages is not well understood. This study focuses on qualitative and quantitative features of full-length TMPRSS2-ERG transcripts in prostate cancer and the sub-cellular localization of the encoded proteins. Cloning and sequencing of full-lengthTMPRSS2-ERG transcripts were performed from a cDNA library generated from pooled RNA of six TMPRSS2-ERG fusion positive prostate tumors. Two types of TMPRSS2-ERG cDNAs were identified: type I, which encodes full-length prototypical ERG protein (ERG1, ERG2, ERG3), and type II, encoding truncated ERG proteins lacking the ETS domain (ERG8 and a new variant,TEPC). In micro dissected prostate tumor cells from prostate tumor specimens increased ratio of type I over type II splice forms showed a trend of correlation with less favorable pathology and outcome. Differential expression ratios of type I/type II ERG transcripts by quantitative analysis correlated with poorly differentiated tumors. Qualitative and quantitative features of specific ERG splice variants defined here promise to enhance the utility of ERG as a biomarker and therapeutic target in prostate cancer.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 May 2008 (29.05.2008)

PCT

(10) International Publication Number
WO 2008/063769 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2007/080826

(22) International Filing Date: 9 October 2007 (09.10.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/850,254 10 October 2006 (10.10.2006) US
60/929,505 29 June 2007 (29.06.2007) US

(71) Applicants and
(72) Inventors: **SRIVASTAVA, Shiv** [US/US]; 13216 Maplecrest Drive, Potomac, MD 20854 (US). **DOBI, Albert** [HU/US]; 7332 Eden Brook Drive #1011, Columbia, MD 21046 (US). **SREENATH, Taduru** [US/US]; 13519 Sanderling Place, Germantown, MD 20874 (US). **PETROVICS, Gyorgy** [US/US]; 9103 Friars Road, Bethesda, MD 20817 (US). **SUN, Chen** [CN/US]; 4901 Battery Lane #203, Bethesda, MD 20814 (US).

(74) Agent: **GARRETT, Arthur, S.**; Finnegan, Henderson, Farabow, Garrett & Dunner, LLP, 901 New York Avenue, NW, Washington, DC 20001-4413 (US).

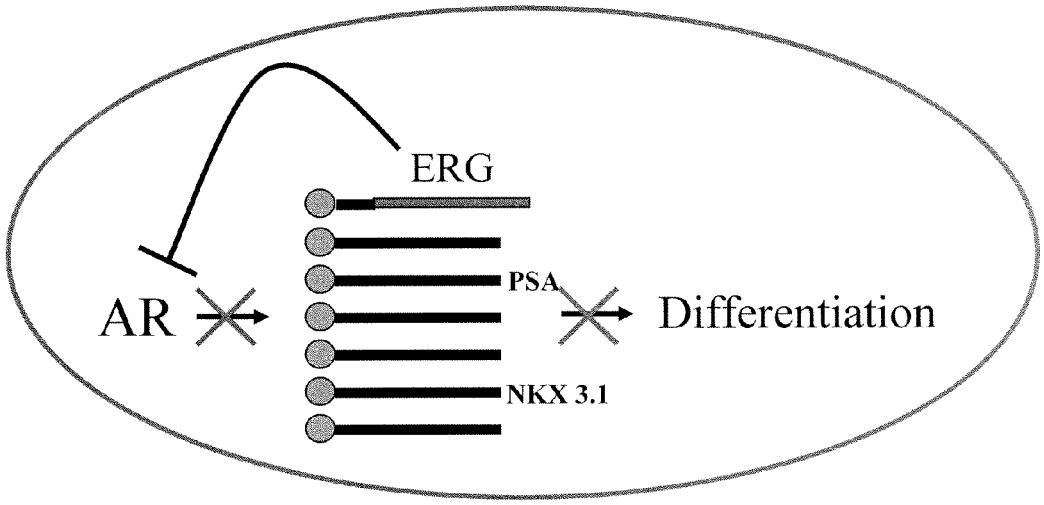
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

[Continued on next page]

(54) **Title:** PROSTATE CANCER-SPECIFIC ALTERATIONS IN ERG GENE EXPRESSION AND DETECTION AND TREATMENT METHODS BASED ON THOSE ALTERATIONS



(57) **Abstract:** The disclosure describes alterations in *ERG* gene expression. ERG isoforms and promoter sequence of the *ERG* gene that are involved in, or associated with, prostate cancer are provided. The disclosure further provides therapeutic compositions and methods of detecting, diagnosing, prognosing, and treating prostate cancer, including biomarkers for detecting the expression of two or more of the following genes: *PSA/KLK3*, *PMEPA1*, *NKX3.1*, *ODC1*, *AMD1*, and *ERG*.

WO 2008/063769 A2